A PRIMER AND GUIDE TO MODELING FOR OPERATORS

GRADUATE RESEARCH PROJECT

Juris L. Jansons, Major, USAF

AFIT/GOS/ENS/05-09

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GOS/ENS/05-09

A PRIMER AND GUIDE TO MODELING FOR OPERATORS

GRADUATE RESEARCH PROJECT

Presented to the Faculty

Department of Operations Research

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Operational Sciences

Juris L. Jansons, BA

Major, USAF

June 2005

AFIT/GOS/ENS/05-09

A PRIMER AND GUIDE TO MODELING FOR OPERATORS

Juris L. Jansons, BA

Major, USAF

Approved:

____________________________________ ________

J. O. Miller, PhD Date
Associate Professor of Operations Research
Department of Operational Sciences

AFIT/GOS/ENS/05-09

Abstract

Operators rely more and more on models to accomplish their work; examples include the weapons employment zone displays in cockpits, logistics models for deployment, and battle simulations to decide courses of action. They often do not have much exposure to modeling, and the products they are using do not always supply adequate documentation. The first portion of this paper serves as a primer on modeling for operators. It then proposes a matrix of questions that an operator should know to ask about any model he is using. The next section contains several examples to illustrate the discussion. The last section includes a proposal to use the matrix as a standard format for modelers to pass relevant information to users. If the operators know which questions to ask, and modelers can embed that information inside the models, then overall effectiveness should increase.

Acknowledgments

I would like to express my thanks to my faculty advisor, Dr. J. O Miller, for the latitude to explore how IDE students can contribute outside of the previous AFIT paradigm. I would also like to thank my fellow IDE Operations Research Students; stay with the herd.

Juris L. Jansons

Table of Contents

Page

List of Figures

List of Tables

A PRIMER AND GUIDE TO MODELING FOR OPERATORS

I. Introduction

Section II of this paper is a primer on modeling for operators since more and more frequently, operators are required to use models. Consider my experience as an F-16 pilot. To plan a mission, I use software models to pick the correct munitions for my target. I use software models to plan my attack and determine if I have enough gas. The information on potential threats relies on models of one form or another. When I fly my mission, my radar warning receiver is a model that characterizes incoming signals to identify who is hostile. When I engage a hostile target, be it on the ground or in the air, my jet uses models to show me where I can launch. Finally, if I have less gas then I would like for the trip home, I can have the jet calculate the ideal altitude and airspeeds for my trip home using an internal model to optimize fuel consumption.

This ever-increasing reliance on models is not likely to reverse itself. In fact, the drive to fuse and synthesize sensor information into simple, intuitive presentations will make us more reliant on the models that feed the pretty pictures. This reliance is not just limited to the cockpit; military decision makers have used models for ages in campaign planning, logistics planning, and systems acquisition. Section III of this paper identifies several questions, in the form of a matrix, that an operator should know to ask about any model he is using to ensure he gets appropriate information. The questions revolve around identifying assumptions. Good operators have always known that you must know the assumptions; this paper is merely an attempt to create a framework to learn and analyze the assumptions within a model. Correct identification of assumptions prevents some of the insidious mistakes that sometimes catch us. Section IV has several examples of issues with models and the associated matrix for each situation.

Increased modeling literacy is only a portion of the battle. That information must be readily available. This paper can also benefit modelers since the same framework that operators can use to learn about models can serve as a mechanism for modelers to communicate with operators. From a model builder perspective, Section V highlights potential misuse of models by operators when assumptions are not clearly communicated. Section VI expands on the matrix discussion to include a few recommendations for model builders. With the exception of the Analysis Capability Flag (ACF), none of the recommendations are new. They are the validation of previous lessons learned via one operator's experience dealing with models. That these recommendations come chiefly from my experience is the largest limitation of this paper. In the language of statistics, it is a sample size of one. I hope that this limitation may also make the paper relevant; it is an operator's perspective. Before embarking on this venture, I interviewed several classmates to find that the issues I address were not unique. I also found through interviews with several modelers that the mistakes generalized in section V have historical validity. Finally, a review of unclassified F-16 Student Weapons School Papers found that many of the papers dealt with effective use of particular models. Those sources that are not directly referenced in the paper are included, following the bibliography.

Because this paper addresses two audiences, it may not completely satisfy either. For operators, some of the material may seem too esoteric or technical. My goal was to keep it simple and relevant, but some technical subjects were unavoidable. For modelers, much of the description may seem to gloss over important details. From both audiences, I ask patience. For operators, if you make it to the end, my hope is that you will find a coherent formulation for dealing with the models that surround you. For the modelers, you should skip to section VI, where the observations of an operator with some exposure to modeling may be relevant, if nothing else, as a demonstration of our frame of reference.

II. The Problem

As an F-16 pilot, there is not a phase of my mission planning or execution that does not in some fashion rely on models. As a neophyte fighter pilot, I had little concern for the limitations of these models, or even conscious awareness of my reliance on them. With more experience, I found several cases where there were issues with my understanding of the model, or the models themselves. As a community, we often relied on subject matter experts, who would delve into the model and find applicable rules of thumb, showing up as weapon school papers or local procedures. Whether we wanted to or not, we were often forced to find general principals and rules of thumb to govern our use of models. We should not have had to develop rules from our experience. If we had a logical framework for thinking about models, we could have done more nuanced and rational assessments. My goal is to translate the existing literature on many of these issues into terms that are relevant to operators, using examples to illustrate potential pitfalls of misunderstanding or misapplication. The problem is not only one of operator literacy, but also communication between operator and modeler. Operators who ask better questions will also need to have the answers to those questions readily available. The same structure for asking questions, can also serve as a means for modelers to communicate assumptions to operators.

III. Modeling Primer

Before embarking on a discussion of modeling, we must place models in the proper context. Models exist only to assist the decision maker. They exist either to analyze problems or to aid in training. For analysis, they provide a tool for understanding the problem; they do not make the decision. Even if they suggest an answer, the decision maker must still weigh whether that recommendation applies to the decision at hand. They are only useful in as much as they aid the decision maker. Ultimately, our focus must remain on the operator, war planner, or commander that makes the life or death decisions. With that said, we can examine one of the tools available to decision makers- the model.

Vocabulary

We must first consider vocabulary. The people who build models have developed their own vernacular that does not necessarily carry over to operator speak. This paper will use terms that are for the most part familiar to operators. However, it will use the word stochastic to describe processes where an element of chance or randomness exists. I never ran across the term in operator life, but found it impossible to read the modeling literature without knowing its definition. We shall also define models as "a physical, mathematical, or otherwise logical representation of a system, entity, phenomenon, or process" (Department of Defense , 1995: A-6). Examples of models include simulators, computer software to plan weapons effects, and simulations of air wars.

There are also potential vocabulary pitfalls for modelers reading this paper. Later when confidence reporting is discussed, this does not imply a specific method of assessing outputs such as confidence intervals. The term is meant to summarize a model user's need to have some relative measure of certainty.

Objectives

Operators are very familiar with objectives. We decide what we want or need to accomplish and this drives and prioritizes our actions. If I am flying a mission to teach a brand new F-16 pilot how to land the airplane, I am not going to spend my time talking about using the radar for air-to-air employment. Likewise, models are built for specific objectives and these objectives will define how the model is structured, what assumptions are made, and to what level the model is abstracted (Law and Kelton, 2000).

Most models of the United States Air Force use computer processing, although arguably constructs like the five rings of Warden or Value Focused Thinking are also models. This paper will mostly focus on the computer models, and as such, it is useful to describe how computer models arrive at an answer, be it the maximum effective range of a missile or assessing the value of opening a second front during a campaign.

Table Look up versus Calculations

Although not a rigorous formal definition, it is useful to put a model into one of two categories, based on how it produces its answers. The model can either use a table look up scheme or do the calculations. Table look up relies on work that has gone before. For example, with mission planning software, the value for how much gas an airplane burns at a given altitude and airspeed is often contained in extensive tables of values that come from flight-testing or very detailed engineering models. The data table exists in the model, or is accessible to the model, and when the model needs the information, it simply looks up the value. Another example is a Radar Warning Receiver that compares incoming signals to stored values to find a match.

The other way that models get answers is to "do the math." These models will have some degree of table look up, but will then calculate the result based on other parameters. The user may have entered these parameters, or they may depend on what has happened earlier within the model. A good example of a calculating model is a missile fly-out simulation. The operator enters the desired altitude and airspeed for the shot. The model contains tables of data on how the missile motor develops thrust, how heavy it is, and what guidance logic it uses. With these values, it then uses physics equations to find the missile's flight path. In these types of high-detail engineering models, it is obvious that all the inputs must be accurate if the result is going to be accurate. When modeling large-scale scenarios, like a many versus many air battle, modelers will often summarize the details of each missile as a table of values. Instead of calculating each missile's fly out to its impact, a simple probability of kill might be assigned to each missile at launch. At the campaign level, the summary may be as general as what the probability is that each type of airplane is victorious against a given opponent. It is important to note that calculation does not have to lead to a specific value (Hartman, 1985a). It could also be a characterization of how things randomly happen. For example, modelers might use a statistical distribution, which fits historical data, to model the rate that F-15s break. The model then draws randomly to see how often an F-15 is down for maintenance (Law and Kelton, 2000; Kelton and others, 2004).

The separation of processes into table look up and doing the math is useful to highlight areas of potential problems. Obviously, the chief source of error in table look up involves erroneous data. If we try to characterize an enemy missile system but our assumed value of its acceleration is incorrect, our model will give us erroneous results. The old adage "garbage in, garbage out" describes this situation. Another potential issue is how the computer deals with an absence of data. When I try to plan a flight for an F-16 at an altitude and airspeed that was not charted, how does the flight planning software deal with that? Does it simply respond that it cannot give me an answer? This is annoying to users, but prevents potential errors from using an inappropriate model. If there are two close data points, the model could interpolate between those two values. Interpolation is viable, but is an approximation. Presumably, this *interpolation* will not cause problems for the designed purpose of the model, but may affect the output if the model is used for an alternative objective. Finally, if the value the model needs falls outside of its tables, does the model use the information it has in an attempt to guess what happens in a region where it has no data? This *extrapolation* runs the risk of entering a realm where the previous relationships do not hold and gross errors are present, such as the difference between subsonic and transonic flight.

Abstraction

The issues in analyzing how a model does its calculations tend to be more subtle because they involve the model's abstractions. To model anything, there is a necessary simplification or abstraction from reality. The objective of the model will drive how the abstraction is accomplished, and the abstractions in turn influence the calculations. Some detail-oriented operators may equate abstraction with loss of truth, but abstraction is not necessarily a negative thing. Well-done abstraction captures the key elements of a problem so that solutions and models can be implemented. Abstraction is only bad if the abstraction fails to support the objectives of the model. The Navy use of anti-air warfare (AAW) models is illustrative. In the late 1950s the first attempts at computer models were "naively simple models (the people weren't naïve, but they had to get on with decisions with limited computer power)" (Hughes, 1997; 26). When sufficient computer power arrived in the 1970's for extremely detailed models, analysts discovered that analyzing the complex results was problematic. They then reversed the trend toward complexity and began isolating the important variables and parameters to arrive at models of "sophisticated simplicity" (Hughes, 1997; 26) As a rule, the need for abstraction, and the potential for misapplication, increases as one moves away from trying to model things and into modeling groups of things or people.

Put another way, abstraction is picking what I care about and dropping the information that is not relevant. The implications of these choices can be very subtle, and are best understood by keeping the model objective in mind. As an example of the interaction of objective and abstraction, consider yourself standing in a furniture store. If your objective is to see if you can convert the store into a warehouse for large equipment, the only information you need are the dimensions of the room. You are able to put everything you need into three sets of numbers. If instead you wish to exit the room, then you need the locations of all the pieces of furniture in the room so that you do not run into them, but their color and texture is immaterial. Finally, if you need to buy a new piece of furniture, then the color and texture of the furniture is important, but not anything about the room.

Model Life Cycle

Models also have a life cycle. Like any system, they get progressively better through refinement. There can be refinements to the user interface, the data tables, or the calculations. Obviously, a well-managed model usually develops better cosmetics and usability for the customer with each generation. The subtler, although often substantial, changes involve either the tabular data or calculations inside the model. While one can find these changes documented in release files, they are not always conveniently visible to model users.

If the model has different systems or entities that it models, each of those may be at a different point in their life cycle. Consider a piece of software that models two missiles. Missile A is well understood. It has been in service for years and extensively tested with the Weapon System Evaluation Program (WSEP). Missile B is brand new. Obviously, the body of knowledge associated with Missile B will not be as extensive as Missile A and therefore the outputs concerning Missile A will be of better quality than Missile B.

Model Scope and Function

The next step towards asking structured questions about a model is to understand some of the formal ways to group models. These groupings, or taxonomies, often reveal important structural elements of an intended model.

A useful taxonomy is to group models by their scope (Department of Defense, 1995: 2-2), sometimes displayed as a pyramid (Miller, 2004). The lowest level contains models that look at a physical system's sub-assemblies. These are also called engineering models. The next level up looks at the system as a whole and how it works during an engagement. Above that is the mission/battle level, where groups of these systems interact and the net outcome is determined. Above that is the theater/campaign level where the battles are further aggregated to see what is happening at the theater level.

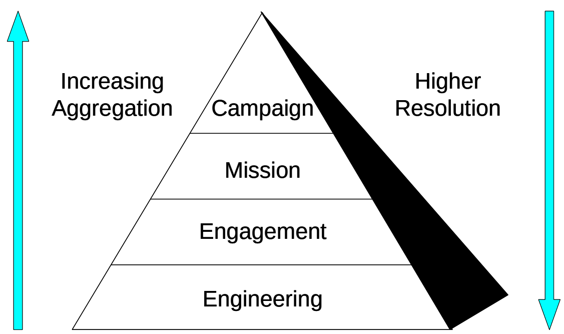

Figure 1. The Modeling and Simulation Pyramid

As we move up the pyramid, the models become progressively more abstract. More things are lumped together and more interactions are summarized with simplifying abstractions. This is necessary and accomplished by using objectives to guide the assumptions and abstractions. What is not modeled should be irrelevant to the objectives of the model, just as advanced radar techniques are not useful to someone learning to land an airplane. If the objective is to see how many bombs to preposition in a theater, then it is critical to model how many bombs the plan requires, not the flight path of each bomb.

Another taxonomy groups models into three functional areas: training, analysis, and acquisition (Department of Defense, 1995: 2-2). Flight simulators are a classic training model. Another training model is having a computer simulate a war, so that a commander and his staff can practice generating the Air Tasking Order (ATO) during a developing situation. If the functional area is training, then the fidelity of the model may be less important. For exercising ATO generation, the results of the air war are less important than the actual process of dealing with changes. Examples of analysis models include; weaponeering software to find the ideal weapon for a given target, computer generated missile fly outs for estimating enemy first launch opportunities, and simulations to aid in war planning. For acquisition decisions, models are made based on assumptions about the future environment to decide what capabilities we need or which missile will better meet our needs. When the purpose of the model is analysis or acquisition, then fidelity is critical.

The taxonomies of scope and function can be summarized as a three dimensional cube. The horizontal slabs show the scope of the models. The higher slabs have greater levels of abstraction. The vertical slices divide each slab into functional areas. Note that this cube can also be sliced into sections for the different services, but this paper will not discuss inter service issues. The cube can provide users with an important first step in understanding their model since, with any model, one can fix where the model is on the cube. The relevance will become clearer as we discuss using models for other than their designed objective. When this happens, the user can visualize crossing lines on the cube and will need to ask a variety of questions, introduced in section IV.

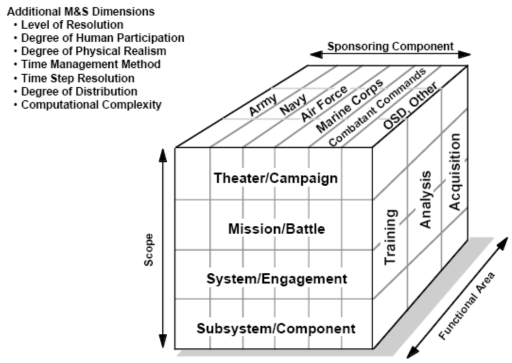

Figure 2. Scope and Function Cube

Figure 2 comes from a Department of Defense (DoD) document giving guidance on Modeling and Simulation (M&S) (Department of Defense , 1995: 2-2). This depiction separates acquisition from analysis, but acquisition is in some ways a subset of analysis. For acquisition, one is simply analyzing the expected requirements to find the most useful system. The remainder of this paper will not explicitly differentiate acquisition from analysis.

Deterministic versus Stochastic Models

Another useful taxonomy identifies if the model is deterministic or random (stochastic) (Hartman, 1985b). In a deterministic model, there is no element of chance. If I shoot a missile with the same parameters each time, it will reach the target in the same way each time. If the first run of the model hits the target, so will all the other runs. In a stochastic model, chance exists. When I shot the missile, just because it hit the target the first time does not mean it will hit it the second time. Whether you want a deterministic or stochastic model depends on what question you want to answer. If you want to know what the maximum launch range of an enemy missile based solely on his hardware, then assuming that all things function as advertised and using a deterministic model will give a worse case answer. If you want to model the expected loss ratios, then you want some real world randomness to include the chance that some missiles fail.

Based on my experience, operators often under appreciate important issues for stochastic models. With stochastic models, they are only useful if you run them multiple times. If I run the model only once, I have only one data point out of all that are possible, so I have no idea of whether that is an average result or not. As an example, I am considering whether to buy a raffle ticket for a one million dollar prize where one million tickets were sold. I build a model based on buying one ticket. I know that since I buy just one ticket, I win with a one in one million chance. I run the model once, and it shows that I win the million dollars. Does that mean that every time I play the lottery I will win a million dollars? If I then run the model 1,000 times, I will see that most of the times I do not win, and the value of that win (or outlier) is slowly averaged out. I see that in reality, I probably will not win. This example seems ludicrous, but that is because we understand the model. With complex stochastic models, such understanding is not readily available, so we can see what the expected outcome is only if we run the model many times.

Another issue with stochastic models is variance. Variance is simply a measure of how far the outcomes deviate from the expected value (Ross, 2003). The impact point of an unguided munition is a good example of variance. When released, each bomb will fall with a slightly different trajectory from all the other bombs of the same type because of minute differences in each weapon (Chamberlain, 2004). The result is a certain amount of unpredictable dispersion, or variance. Any model that includes randomness will also have variance. Whether the model presents this information will vary from case to case, but should be something operators take the time to find and understand.

Stochastic models are developed to simulate the randomness and variability of real life, but there is no guarantee the modeled quantities are equivalent to real life. Models are by necessity simplifications of the real world and these simplifications may limit or exacerbate the qualities of randomness and variance (Miller, 2004). Assume that in the mid 1990s, I made a model of the stock market, where my model assumed that the market had a 50% chance of going up at 3% over three years and a 50% chance of going down the same amount. If I had invested, the dot com boom would have created significantly better performance than my model had predicted because life was more variable than I had modeled.

Descriptive versus Prescriptive Models

A final useful taxonomy segregates models in descriptive versus prescriptive types (Hartman, 1985b). A descriptive model “describes how a system will operate if values for all the input variables and decision rules are given by the model user” (Hartman, 1985b:1-3). For example, planners can use a simulation of the opening phases of a war plan as a descriptive model. They observe the results of the simulation and then make inferences about what is going on and what the best course of action is. Another example is a weaponeering program where operators try out different munitions against a given target to find which ones give the greatest chance of achieving a kill. Contrast this type of model with a prescriptive model that "specifies how the system ought to operate to achieve some objective" (Hartman, 1985b: 1-3). With prescriptive models, users usually input the parameters of the problem, and the model provides the answer. For a weaponeering program, operators would input what type of target they wanted destroyed and the model would tell them which munition to use. The key issue with prescriptive models is that they must make enough abstractions to make the problem solvable. Whether this is permissible without loss of applicability is up to the decision maker. Further, since the prescriptive models show you *the* answer for your assumptions, sometimes they do not provide information on other alternatives. It is often useful for prescriptive models to show a range of choices, since the decision maker may have other factors outside the purview of the model that make second or third choices the best ones.

Sensitivity Analysis and Confidence Reporting

Sensitivity analysis is varying the range of an input or assumed values to see what happens to the output values (Clemens and Terence, 2001). Since information drives models, sensitivity analysis provides a means to check how important the various inputs to the model are. Some inputs may be more critical than others. If battery life is the chief limitation on a missile's maximum range, then seeing how much the launch range increases with a small increase in battery life provides an indication of how important that parameter is, and more importantly the potential consequences if the value is wrong.

Lastly, some measure of the confidence of the output is helpful. It depends on the model, but often models do not show this data or leave it buried inside the documentation. For example, equipment used to identify contacts as hostile or friendly usually has some percentage of confidence. Often operators must study the accompanying documentation to internalize these confidence measures, whereas they could be graphically depicted during operations. That is not to say that it is never visible, some data links have such displays, but often there are opportunities to make this data more readily available. Operators must also understand that there are a variety of ways to quantify and describe confidence, and the exact implementation will depend on the model.

Model Verification, Validation and Accreditation

The Department of Defense (DoD) has actually recognized the increasing reliance on models and attempted to guide and standardize their use. The joint organization for modeling is the Defense Modeling and Simulation Office (DMSO). Each service also has organizations responsible for modeling and simulation (M&S). These can found at www.dmso.mil. One of the key responsibilities of these organizations is guiding Verification, Validation and Accreditation (VV&A or sometimes VVA of models). Verification is "the process of determining that a model implementation accurately represents the developer's conceptual description and specifications" (Department of Defense , 1995: A-8). In short, verification ensures that the model is "built right." Validation is "the process of determining the degree to which a model is an accurate representation of the real-world from the perspective of the intended uses of the model" (Department of Defense , 1995: A-8). In essence, validation ensures that the right model was built. Accreditation is "the official certification that a model or simulation is acceptable for use for a specific purpose" (Department of Defense , 1995: A-8). The

DMSO and service organizations are important steps in the process of quality control and interoperability in modeling.

Properly implemented VV&A processes ensure that models have met the requirements of end users (Law and Kelton, 2000). However, even if a model has been through the VV&A process, operators can still create problems by misapplying the model. Nor will all models necessarily be covered by VV&A. For example, does the weapons envelope zone display within a cockpit fall under the VV&A umbrella? There is also the potential that the model, or data therein, goes out of date. While VV&A is intended to be a continuous process, assessing whether the VV&A process of the USAF adequately tracks and guides this development is beyond the scope of this paper.

With the taxonomies and issues previously discussed, users should be in a position to examine each model at their disposal. What follows is a structure for identifying the salient points of a model. To borrow the concept from a popular series of books, what follows is the idiot’s guide.

IV. Idiot's Guide

Reminders

The Idiot's guide is written from the operator's perspective. The guide is intended to focus the model user's efforts to understand a given model. Modelers should understand upfront that this is not an extensive categorization of all possible issues, but is designed instead to highlight potential problems.

Models are very useful tools, but like all things have their limitations. An oft-cited dictum is that, "all models are wrong, but some models are useful" (George E. P. Box). To make them useful, we should understand them; the matrix below provides a framework to develop our understanding of each model. The matrix draws on the vocabulary and concepts established in the previous section. The left most column are the areas of interests such as scope and function. The middle column covers how the areas of interest apply to the model user. It is labeled "about me" to highlight that these questions are from the viewpoint of the operator. This entire section will maintain this perspective. The far right column pertains to the model. A difference between the user and model in an area of interest highlights what the operator must further examine.

Below the matrix is commentary on each of the cells, associated with the cells by the number in the title. The first digit of the number is one for the user, two for the model, and three for the relationship between the two. The digit after the decimal refers to which particular area of interest is involved. The discussion of the relationship between the user and the model is not exclusively limited to that area of interest. The areas of interest merely provide a logical starting point for the discussion.

The Matrix

Table 1. The Matrix

Areas of Interest	1.0 About Me	2.0 About the Model
X.1 Objective	1.1 What is my question?	2.1 For what purpose was the model designed?
X.2 Functional Area	1.2 What is the functional area of my question?	2.2 What functional area is the model in?
X.3 Scope	1.3 What is the scope of my question?	2.3 What is the scope of the model?
X.4 Deterministic or Stochastic	1.4 Do I need deterministic or stochastic?	2.4 Is the model deterministic or stochastic?
X.5 Data sources and Calculation	1.5 What is the output accuracy I require?	2.5 Does the model exclusively use table look up or do calculations as well?
X.6 Descriptive or Prescriptive	1.6 Do I need the model to show me what action to take or to describe the problem?	2.6 Is the model prescriptive or descriptive?
X.7 Chief Abstractions	1.7 What type of abstractions am I willing to accept?	2.7 What are the chief abstractions in the model?
X.8 Sensitivity Analysis	1.8 Do I need sensitivity analysis?	2.8 What capabilities does the model have for sensitivity analysis?
X.9 Confidence Reporting	1.9 What sort of confidence reporting do I need?	2.9 What capabilities does the model have for confidence reporting?

1.0 About Me

1.1 What is my question?

Obviously, I have to know what my own objectives are before I decide whether a model will benefit me.

1.2 What is the functional area I am looking at?

I must decide whether my goal is to analyze a problem, or train a skill set.

1.3 What is the scope of my question?

I need to define if my question relates to the subsystem, system, engagement, mission, or campaign level.

1.4 Do I need deterministic or stochastic?

This is not usually something that I am directly concerned with. Modelers will usually make the model deterministic or stochastic based on how they are solving the problem. There are maybe two times when it is a driving concern. If I need repeatability, like in the simulator, then I probably want a deterministic model. A deterministic model is also useful if I want to look at an absolute worst or best case, like missile fly-out. I need a stochastic model if I am looking for unpredictability. It is best that a battle simulation for campaign planning has random performance to capture a sense of the unpredictability of the real world.

1.5 What is the output accuracy tolerance required?

I must define the precision I need from the model. For the engagement models and below, it is possible to get outputs that predict absolute performance. The limits on output precision will come from the assumptions and calculations. For physical systems that are well understood, it is reasonable to expect higher tolerance. If I have used thousands of 2,000 lb bombs, then I can probably model them very well, within the constraints of their inherent variability.

The higher levels of the cube demand more assumptions and abstractions, so I must be less concerned with precise values and more with qualitative and relative values.

Say we model two aircraft fighting each other within visual range to evaluate the effectiveness of high off-boresight (HOBS) weapons. The model says aircraft equipped with HOBS weapons win at a two-to-one ratio. We cannot say that in combat we expect a two-to-one exchange ratio. Outputs at this level really only carry relative weight. What we can say is that generally HOBS weapons are very helpful and place the non HOBS equipped aircraft at an extreme disadvantage.

1.6 Do I need the model to prescribe what action to take or to describe the problem?

This question is really only relevant to the analytic model. If I want the model to tell me what it thinks I should do, to prescribe a course of action, then the model must have a logical means of solving the problem. Yet, turning a real world problem into a logical problem may drive up the number and level of abstractions present in the model. Consider for example a technique that optimizes where to place a new facility in relation to its customers. I tell the model where all the customers are, and then it returns the precise location that is central to all of them. The model assumes that all locations are possible, and so it may return an unavailable location. If my customers are in Japan and California, it may say put the new facility in the middle of the ocean. The model does not know that this location is not possible. That is not to say this model is not worthwhile. It shows the “ideal” location based on what it knows, and I can then modify the real solution and chose Hawaii as a logical solution. Prescriptive models are especially helpful and applicable in cases where the chief issues lie in dealing with a large amount of data. If I do not need a prescriptive model, then by default I am looking for a descriptive model.

1.7 What type of abstractions am I willing to accept?

This paper focuses on using existing models, not designing models. So the better starting point for this question will be with the chief abstractions in the model (2.7). The operator must examine these abstractions and then relate them back to the situation to see if they are acceptable for his case.

1.8 Do I need sensitivity analysis?

When I am unsure of the data that I am entering, it would be nice to run sensitivity analysis to see just how accurate my data needs to be to avoid significant changes in my model output. It is possible to confuse sensitivity analysis with the need to input a range of values to find an answer. Putting in a range of values to examine a problem technically falls under the category of "design of experiment (DOE)." DOE is where the analyst is interested in the performance of a system over a well defined set of values for specific factors. When operators are asking these sorts of questions of models, our design of experiment is usually to pick a set of values for an employment option that we are interested in and manually crunch them. For example, if I plan on diving at 20 degrees to deliver my bomb, then I may run the 15 and 25 degree numbers to see how my minimum release altitude changes. For the more abstract problems, like finding out if a there is a particular range and azimuth where a radar would have a sweet spot, operators should rely on trained analysts to design the most efficient test profiles.

1.9 What sort of confidence reporting do I need?

This question is related to what sort of output accuracy is desired, but focuses on the quality of the output. There is no formally defined term "confidence reporting." The term is my attempt to lump a variety of measures of confidence into one category with a plain English title. Thus, for this paper, confidence reporting means any measure of merit for the output. For example, it could include color flagging of where the model is in its life cycle, discussed in Section VI. It could be a cockpit indication that toggles back and forth between two different values. For stochastic models, it could be formal statistical tests for differences or confidence intervals. My goal is not to discuss at length all the different ways to measure confidence, but just to point out the importance of this measure. Operators should be aware that the confidence measures available will depend on how the model is implemented.

2.0 About the model

2.1 For what purpose was the model designed?

Every model is designed to support some objective. It is important to know the designed purpose of a model to see if it will support your situation. Ideally, users should only use models which are designed to support their objective. With many models, the objective is obvious. A radar warning receiver is designed to identify if other radars are tracking the aircraft and pass that information to the operator. There are however many cases where the objective is not so obvious. For example, a battle level model may depict the air space over Kosovo. It might be tempting to use it to see how introducing a new weapon would have influenced the fight. That would be a mistake because this model was designed to examine the importance of communications links with the Combined Air Operations Center (CAOC) for time sensitive targeting, and thus may not have the weapons fidelity required for your purpose.

Related to the purpose of the model is determining who made the model. Different organizations and services use different assumptions. A model designed to train certain pieces of doctrine may underemphasize others (Hughes, 1997: 48)

2.2 What functional area is the model in?

Was the model designed to train skills or analyze problems?

2.3 What is the scope of the model?

Determining the scope of the model will suggest its level of abstraction. A theater/campaign model will have the most amount of abstraction. Individuals will be grouped into units and the results from these aggregate entities will be very general. Moving down the cube to the mission or battle level models will increase the detail level, but there will still be some sort of aggregation. At the next level down, the system or engagement level, there will be even more detail, until finally at the subsystem or component level, I will have the greatest detail (Friel, 1997: 141-162)

2.4 Is it deterministic or stochastic?

Hopefully, the model documentation includes whether it is deterministic or stochastic. If I am trying to determine if it is deterministic or stochastic based solely on output, it can be difficult. Deterministic models will give the same answer each time the same input conditions are used. For stochastic models, the same input conditions may or may not lead to different outputs.

Stochastic models need random numbers to create random events. The discussion of random number generators is beyond the scope of this paper, but a useful way to view them is as a really, really long list of numbers that have no relationship to each other (Law, A. M., and W. D. Kelton, 2000). When a stochastic model starts, it picks a place on the list and proceeds from there. Modelers can usually specify whether or not the model starts in the same place on the list each time. Thus, if I rerun a stochastic model with the exact same set up and the random numbers start at the same place, the results will be the same as my first run.

2.5 Does the model exclusively use table look up or do calculations as well?

If the model is only table look up, the currency of the tables is of chief concern. The other issues may be the precision of the tables and their extensiveness. If there are holes in the data, how will it handle interpolation or extrapolation? If the model does calculations, then not only should I be concerned with the currency of the tables, but also with how well the model does the math.

2.6 Does the model describe or prescribe?

If I tell the model what my objective is and it tells me what the answer is to achieve my objective, then it is a prescriptive model. If I put my information into the model and then have to examine the output to decide what to do, then it is descriptive.

2.7 What are the chief abstractions in the model?

For models higher up on the cube, this is usually the hardest element to find and understand. Even if the assumptions and abstractions are identified up front, it can be difficult to understand the significance of those abstractions. A variety of techniques, some with cryptic names, may appear. Examples are linear modeling, Bayesian networks, and Markov chains. The best way to understand these techniques is to discuss their significance with the model builders to identify if those abstractions will interfere with your use of the model.

2.8 What capabilities does the model have for sensitivity analysis?

This question is really only pertinent to analysis work. The techniques available will depend on how the model was created.

2.9 What capabilities does the model have for confidence reporting?

As previously mentioned, the sort of confidence reporting available will depend on how the model is constructed. There are a variety of techniques, and hopefully the presentation is simple and understandable.

3.0 The model and me, or can the model handle my question?

3.1 Objectives

The first sanity check is to make sure that the stated objectives are somewhat similar. With some common ground, I can turn to the remaining questions.

3.2 Am I changing functions?

If the model's functional area is different than what I need, then I must be very careful. It follows that the model was built on different assumptions than I would have made. I will likely find potential problems in the calculations or abstractions of the model.

Moving a training model into the analytic realm is usually the most dangerous. The training model may have sacrificed accuracy for ease of operability, and this may cause problems (see examples 1 and 2).

If I am moving from an analytic to a training function, then the model can probably be useful in its new role. I should still examine the assumptions, but usually analytic work requires sufficient details that it can also work as a training model. A possible exception is that analytic work may often involve stochastic models. If, for my training needs, I need to reward correct task accomplishment with specific results, this model may not be effective.

3.3 Am I changing scope?

If my purpose is higher up on the cube than the model's, I will have to extrapolate from the limited sample size of the model. It then becomes very probable that other dynamics may enter the picture. If in my engagement level model, the opponent has High Off-Boresight (HOBS) weapons and the one versus one engagement shows that I lose twice as often, I should not extrapolate this into a campaign model and assume I will lose twice as many airplanes as the enemy. In this case, the advantage of long-range weapons and command and control architecture were not modeled in the one versus one.

This type of error is like using an upgrading pilot's performance during basic fighter maneuvers to describe how he will do as a mission commander.

If my purpose is lower on the cube than the model's, then the value that I am drawing from may be overly simplified and not representative. This type of change occurs when people unfamiliar with a capability try to use the model to learn or make judgments. For example, if the campaign level model includes the superior command and control architecture and long-range weapons to summarize that blue beats red twice as often, it would be erroneous to say that any time a red aircraft faces blue, red will lose at this rate. These types of conclusions may also not accurately capture the variability in the data.

3.4 What is the output telling me, deterministic vis-à-vis stochastic?

Models at the top level of the cube are usually designed to include a certain amount of uncertainty. That is not to say there is no deterministic behavior, often the sub processes like losses to enemy action are deterministically modeled. The overarching goal though is to capture the unpredictable nature of interaction of a large number of factors. As mentioned earlier, using one data point to understand a raffle is the same idea as running a simulated air campaign one time and declaring that its results represent what will happen. If one is going to do analysis, it becomes necessary to run several iterations to find what the average outcome is. For training purposes, this is generally not an issue. Consider a command and control exercise to generate Air Tasking Orders (ATO). The planners generate an ATO and then a campaign simulation takes the ATO and finds the day's simulated results. The actual results of the day are immaterial, what is important is that the planners have to execute the next ATO cycle with a slightly different problem.

Often times, the bottom level of the cube is deterministic. In the analytic models the question is usually to find out how a systems works or to look at worst-case scenarios. It may also be that the physical laws being modeled are deterministic, e.g. Newtonian physics. That is not to say that all models at the lower levels are exclusively deterministic. For example, representative random errors could be built into a missile fly out to give a sense for what the probability of intercept is. The operator could then run several iterations to capture the overall probability of kill (P_k). The training models at the lower level are usually deterministic so that the same action will produce the same results. For example, in an emergency procedures simulator, it is important to reinforce that if the crew accomplishes the critical actions in a timely fashion then they can save the aircraft. If the time between a fire light and engine failure varied between 1 and 30 seconds, sometimes the correct procedures would not avert disaster. Once again, not all training models at this level are exclusively deterministic. During simulated air-to-air engagements with opponents, it may be beneficial for the simulator to have missiles kill opponents only a portion of the time.

3.5 Are there limitations in the calculations of the model that make it unsuitable for the use I am envisioning?

Comparing my required accuracy with that of the model is not a precise science, but some general rules apply. The higher up the cube I move, the more likely it is that some sort of table look up is going on. That means that the currency of the tables can be a critical issue. There is usually a lag time for information inclusion in the higher level models. Thus, a radical new assessment of a threat missile may not actually affect the campaign model for several years.

At the higher levels of the cube, the outputs of the model become more a measure of relative merit and less a measure of absolute performance. Thus, it makes little sense to take an engagement level model and input my planned ATO to predict the results. It would be more suitable to examine whether the use of a first wave of stealth attacks might be more effective on average (after numerous runs) than just going downtown with conventional aircraft.

At the lower levels of the cube, the actual value that the model returns is pertinent. That value has physical meaning. How precise it is depends on the nature of the question, but also the assumptions and calculations. If my question requires a certain level of precision, how do I know if the model produces that precision? For example, knowing the maximum range that an enemy can launch is a critical number. Addressing how well the model calculates that is usually beyond my technical means. Section VI addresses this area, since it primarily focuses on better transparency and detailing of confidence reporting by the modelers.

Similarly, I need to have an awareness of when the model extrapolates. Hopefully, my model footnotes or highlights these cases. Interpolation should be sufficiently precise for the models intended use, but if I wish to use it otherwise, I may run into problems.

3.6 Will a prescriptive model work descriptively?

A potential danger exists when a prescriptive model fails to report enough information to function as a descriptive model. If the prescriptive answer is relevant and acceptable, this is not an issue. If however the decision maker does not believe that the prescriptive answer is appropriate than there must be enough information for the decision maker to still use the model, or at least to identify why the model is returning a different answer. For example, there are initiatives underway to provide prescriptive tools for campaign planning (Caroli and others: 2004; Wentz and Wagenhals: 2004). If I am using software to find what target to strike in a terrorist network, and it returns only one target, without showing how the factors interplay based on its assumptions, then I cannot truly weigh the output.

With a descriptive model, the operator has to find the decision. Assuming that a model is designed sufficiently for its objective, it will support my decision. It will not fail in giving me the insight I need to make the decision. Thus the case where I have a descriptive model and think I need a prescriptive one, is really a case where the model is not sufficiently assisting me in identify what are the critical parameters.

3.7 Do any of the chief abstractions in the model make it unsuitable for my envisioned use?

With what I know about the design of the model, do I believe the model can answer my question? If I am not certain, then I need to contact the model point of contact and get an answer to my question. All models involve certain abstractions that are a function of the model objective. By understanding the abstractions, I can asses how to use a model's output. Ideally, I should not have to hunt for these descriptions, and section VI suggests a format developers could use to make these standardized and accessible. With higher-level models, I must also remember that they require more abstraction; I must ensure that my questions tend towards relative value and not absolute solutions.

3.8 What do I learn from sensitivity analysis?

If there is no built-in sensitivity analysis, then depending on the model, I will have to caveat my results. For example, at the bottom level of the cube, if I do not know what a small change in battery life of a missile changes, then I cannot understand if that value really matters for its maximum range. If I do not have sensitivity analysis at the higher level, then I have no way of knowing what my most critical assumptions are.

3.9 What do I learn from confidence reporting?

If confidence reporting is available, it gives me some idea of how certain that model is about its result. Imagine an automatic target recognition (ATR) system that operates through my infrared sensor. If it reports a T-72 tank, I would like to know how certain it is that it is correct. Confidence reporting could also take place prior to the sortie, where I have the option to hide the ATR assessment if it is below a certain value.

For stochastic processes, I must examine the range of values, their variability, and the applicable statistical tests. The range of values may be important since I cannot have values above or below a certain value. Likewise, variability gives me some idea of how erratic the behavior is. Finally, statistical tests have rigorous mathematics behind them, but I can only use them if I understand their significance. In the absence of that, my hope is that the modeler has prepared a simple scheme to convey that information.

V. Examples

The following examples are fictionalized accounts drawn from my experience, interviews with operators and modelers, and emerging capabilities. The first part of each example is an overview of the situation and its resolution. The matrix follows with yellow shading on areas of the matrix that point towards problems. These areas are then discussed, illustrating potential uses of the matrix.

1. B-52 (Change in Function)

You are the Supervisor of Flying. A B-52 has just taken off and lost an engine nacelle. You believe this is the first time this has ever happened and, not surprisingly, there is no specific checklist guidance. You start recruiting help, and the Squadron's Operations Officer suggests that he send a couple of Instructor Pilots over to the simulator. They can set up the failure in the simulator and see if it is possible to land in that configuration.

You also call Boeing about the problem. As the B-52 burns down gas, you get your first report from the simulator pilots who say that it is not possible to land. After a while, you get a report from a Boeing engineer who has found old tech data that covers loss of an engine nacelle. According to the engineer, the pilots should be able to land the airplane. So whom do you trust?

In this case, the pilots use the engineer's statements and attempt a landing. The landing is successful, but you are left to wonder why the simulator was wrong.

Table 2: B-52 Example (Change of Function)

Areas of Interest	About Me	About the Model
X.1 Objective	I want to analyze if it is possible to land the B-52 in a new configuration.	The model is designed to train aircrew in flight procedures.
X.2 Functional Area	I am looking to analyze a problem.	The model is designed as a training model.
X.3 Scope	The scope of the problem is a systems level issue	It is a system level model.
X.4 Deterministic or Stochastic	I need a deterministic model.	It is a deterministic model.
X.5 Data sources and Calculation	I need very accurate tolerance.	It probably has both table look up and do the math.
X.6 Descriptive or Prescriptive	I need model to describe problem.	The model will hopefully describe what happens, so the simulator pilots can pass recommendations.
X.7 Chief Abstractions	I cannot accept any abstractions that will degrade the accuracy of the output.	The model will have some simplifications, although initially I do not know what they are.
X.8 Sensitivity Analysis	I need to know how sensitive the configuration is to different weights, wing damage, cross wind, etc.	The only sensitivity analysis available is to vary the configurations manually and see what changes.
X.9 Confidence Reporting	I do not expect confidence reporting since the model was designed for training.	It is a training model, so I have no confidence reporting available.

The chief disagreement between the operator and the model occurs in the functional area. You wanted to analyze the problem, but the simulator was built as a training model. As is often the case, the training model contains simplified data for the regimes that it trains aircrew. The model abstractions are built for a training purpose, which does not require high fidelity. The data on loss of a nacelle was considered so remote that it was not even considered for purchase and development in the simulator.

2. KC-10 (Change in Function)

KC-10 simulator operators notice that if they dial up crosswinds they can create a condition that is not covered in the performance manuals. Performance manuals discuss wing engine failure during take off with lighter weight conditions. In this case, aircrew face a condition where they have the thrust to take off, but the adverse yaw means that they cannot maintain sufficient directional control. For their abort decision, they use the airspeed required to maintain directional control. In the simulator, with high crosswinds the value to maintain control is 10 knots higher, corresponding to the airspeed at which the airplane could rotate to a two-point attitude and still take off.

Based on this observation, the conservative approach is followed, and procedures changed to use the higher airspeed. Since the performance manuals do not include calculations for this airspeed, large amounts of man-hours are spent to create new data.

Table 3. KC-10 (Change in Function)

Areas of Interest	About Me	About the Model
X.1 Objective	I want to analyze take off conditions with high crosswind and the loss of a wing engine.	The model is designed to train aircrew.
X.2 Functional Area	I need to do analysis	It is a training model.
X.3 Scope	I have a system level problem.	It is a system level model.
X.4 Deterministic or Stochastic	I need a deterministic model.	The model is deterministic.
X.5 Data sources and Calculation	I need very accurate output.	The model will do the math based on tables of data. I do not know how extensive or current tables are.
X.6 Descriptive or Prescriptive	The model should describe the problem.	The model is descriptive: I can see what happens based on my actions, but it does not tell me what I should do.
X.7 Chief Abstractions	I cannot accept abstractions that create artificial performance.	The model probably has a reduced fidelity level to make it more manageable or less expensive.
X.8 Sensitivity Analysis	Sensitivity Analysis would be nice.	Sensitivity analysis is not available. The model was not designed for analysis. Manual runs with different configurations may help.
X.9 Confidence Reporting	I do not expect confidence reporting since the model was designed for training.	It is a training model so I do not expect any confidence reporting capability.

Like in the first example, a look at the matrix shows that there is a change in functional area. The simulator is being used for analysis, and it may not be designed or capable of modeling this situation. After a call to Boeing, the contractor states that the lower airspeeds are probably not necessary, but will not cause any problems. Why the difference? There is a change in function: just like before the simulator is being used to analyze instead of to train and the data for the takeoff is not accurate.

3. Logistics Planning (Stochastic Process)

You are examining how long it will take to deploy a brigade into Korea for a major exercise. You have a desktop tool that describes itself as a stochastic logistics-planning tool. It seems like it covers your objective, functional area, and scope, so you run it and come up with a value. It is a very detailed model and you spend a great deal of time getting the values correct. Because of this level of detail, you are confident that you have captured most of the significant issues and that none of the model abstractions will be an issue for you. You find it will take three weeks to deploy the unit.

After the exercise, you find that you significantly underestimated the time that it would take to deploy the brigade.

Table 4. Logistics Planning (Stochastic Process)

Areas of Interest	About Me	About the Model
X.1 Objective	Determine how long to move supplies into theater.	The model will calculate the time to move supplies into a given theater.
X.2 Functional Area	I need to do analysis.	The model is built for analysis.
X.3 Scope	I have a campaign level problem.	It is a campaign level model.
X.4 Deterministic or Stochastic	I want a stochastic model to capture the uncertainty of real life.	The model is stochastic.
X.5 Data sources and Calculation	I need high accuracy.	The model does the calculations and uses tabular data about aircraft capacities.
X.6 Descriptive or Prescriptive	I want a descriptive model to understand the problem.	It is a descriptive model because it shows what happens. It does not tell me how I am supposed to move things.
X.7 Chief Abstractions	I can accept abstractions that do not affect accuracy of the altitude calculations.	The model makes simplifications about loading schemes, and aircraft configurations, but these are well documented and are not significant to you.
X.8 Sensitivity Analysis	I am not paying attention to sensitivity analysis.	Sensitivity analysis is not available.
X.9 Confidence Reporting	I did not consider this before the exercise.	The model has no built in confidence reporting mechanism.

Should you chalk this one up to the fact that no model will predict the future? That is true, but you could have gotten much better insight if you had run the model a variety of times. Your error was that you did not understand the implications of a stochastic model. Your first run happened to be optimistic. In reality, you need to do more runs to find a reasonably precise measure of the expected performance, or mean. Since this is a stochastic model, you also need to examine the variability of the results. With several runs, you might see that the process is quite erratic, and that your values range between three and ten weeks.

These types of models are also used to “what if” different scenarios to find whether one scenario is better. For example, maybe you wonder whether you should send units via sealift. In this case, you will have to run each scenario multiple times to get a good idea of its mean and variability. Armed with information about each scenario, you can then examine if the average times varied. Simply seeing a difference in values does not mean that one way is better. Say for instance, you ran both scenarios three times and compared results to find that the scenario with sealift took longer. There is always a chance that the runs you did with sealift just happened to be unlucky, and all had long values. Obviously, the more you run each scenario the less likely that this will occur. There are rigorous mathematical tools available to identify significance. If they are not built into the model, or understandable, most likely the model point of contact can help. Other avenues for support are to find a USAF analyst or call the Department of Operational Sciences, Air Force Institute of Technology.

4. Attack Planning (Calculation Errors)

You are a pilot planning your roll-in to drop bombs from high altitude. You use weapons planning software to determine how much altitude you will lose during your five seconds of planned dive. You add that number to the minimum altitude for safe recovery to avoid the fragmentation pattern, thereby calculating a roll in altitude that gives you five seconds of track time. After repeated attacks on the range, you find that your planned 5 seconds of track time is always 3 seconds. After talking with software engineers, you discover that the math calculations do not include the altitude you lose as you roll into the attack, robbing you of 2 seconds of track time.

Table 5. Attack Planning (Calculation Errors)

Areas of Interest	About Me	About the Model
X.1 Objective	Calculate minimum altitude for release of weapons.	The model calculates minimum altitude for release of weapons.
X.2 Functional Area	I need to do analysis.	The model is designed for analysis.
X.3 Scope	I have a subsystem level problem.	It is a subsystem level model.
X.4 Deterministic or Stochastic	I need a deterministic model.	It is a deterministic model.
X.5 Data sources and Calculation	I need accuracy sufficient for safe bomb release.	The model does the math with my input parameters.
X.6 Descriptive or Prescriptive	I can use descriptive model, although a prescriptive for my specific use would not be a problem.	I can use the model to explore a variety of different assumptions, but it is up to me to pick which option is best.
X.7 Chief Abstractions	I can accept anything that does not affect accuracy of altitude calculations.	I can guess that there are abstractions in the math, but do not initially know them. Since the model and your objective are the same, you presume they are not a factor.
X.8 Sensitivity Analysis	Sensitivity analysis is required.	I can get manual sensitivity analysis with multiple iterations.
X.9 Confidence Reporting	Confidence reporting would be great to have.	No confidence information is available.

The first four lines on the matrix match. The error exists within the data sources and calculations.

5. Examining Air to Air Missiles (Data Sources and Model Life Cycle)

You are preparing to face a rapidly improving adversary in air-to-air combat. For several years, he has trained with missile A, but he has just acquired a new missile, B. You find a chart with the launch regions for both missiles. Missile fly out models produced the chart and you concentrate your tactics around its values.

After several successful combat sorties, it seems that the chart values for missile A are much more representative than for missile B. The chart seems to overstate missile B's capabilities.

Table 6. Examining Air-to-Air Missiles (Data Sources)

Areas of Interest	About Me	About the Model
X.1 Objective	I want to know the first launch range for missile A and B.	Calculate first launch range for missile A and B
X.2 Functional Area	I need to do analysis.	The model is designed for analysis.
X.3 Scope	I am looking at a subsystem issue.	It is a subsystem level model.
X.4 Deterministic or Stochastic	I need a deterministic model.	It is a deterministic model.
X.5 Data sources and Calculation	I need high accuracy.	The model strives for high accuracy, and the laws of physics that describe a missile in flight are well understood. The chief issue is whether the assumptions about each missile are good.
X.6 Descriptive or Prescriptive	I want a descriptive model, so I can understand the opponent's capabilities and develop tactics.	The model provides descriptive outputs of range, time of flight, whether it is a hit, etc.
X.7 Chief Abstractions	I cannot accept anything that exceeds my required accuracy.	Many abstractions may be present in the math, but I suspect since it is an engineering level model designed for my purpose that they will not amount to much.
X.8 Sensitivity Analysis	It would be good to see how important the assumptions are for each missile, or at least have some idea of the confidence the modelers place in each result.	The chart does not reproduce any information on sensitivity analysis.
X.9 Confidence Reporting	Confidence reporting would be great to have.	The chart does not report any confidence information.

In this case, the calculations are correct, but the discrepancy comes in the tables of values. This data includes things like how long the battery inside the missile will last, how much thrust it develops, and what its guidance laws are. The model uses this data and crunches away to give you an answer. For missile A, it turns out that the table look up values are very close to its true capabilities. For missile B, the analysts are less sure of its basic parameters. They have chosen conservative assumptions that overstate its true capabilities. This is an example of the life cycle of a model, where the values get progressively better the more we understand a system.

While this example is in terms of a threat system, it is also directly related to blue systems where we can continually strive to improve our understanding of the weapon through live fire events like WSEP. Section VI addresses how this information could be flagged for users.

6. Weaponeering (Sensitivity Analysis)

You and your wingman are preparing to destroy the inhabitants of Saddam Hussein's buried bunker and are using weaponeering software to determine how to set your fuse. The bomb and fuse have to drive through 10 feet of dirt and then a variety of internal structures to explode at a specific depth. There may be some severe limitations on inputs though since your intelligence on the bunker may not be so good. What if the dirt is 30 feet deep? What if the concrete is stronger than you expect? At this point, a sensitivity analysis tool would be very helpful. It could find the range of values where the solution does not change. If it shows that the initial assumptions are too critical, maybe spending the next 4 hours in front of the computer will not produce better results.

Table 7. Weaponeering (Sensitivity Analysis)

Areas of Interest	About Me	About the Model
X.1 Objective	What is the fuse setting so the bombs explode at the correct place?	Determine how deeply a weapon with a given fusing will penetrate
X.2 Functional Area	I need to do analysis.	The model is designed for analysis.
X.3 Scope	I need to look at the subsystem level.	The model looks at the subsystem level.
X.4 Deterministic or Stochastic	I need a deterministic model to find a predicted value. I would like some stochastic measures to capture reliability data.	It is a deterministic model.
X.5 Data sources and Calculation	I need high output accuracy to ensure it explodes where planned.	The model does the math by using input parameters and values in table to come up with answer.
X.6 Descriptive or Prescriptive	I need to describe so I can find answer, although if it could prescribe that would be fine.	This particular model is descriptive. It does not specify the answer based on how much dirt and concrete there is, but instead you must see if a particular setting works for the conditions.
X.7 Chief Abstractions	My only concern is once the bomb hits the dirt, so any assumptions that do not affect that are all right.	I am unsure of the chief abstractions, but suspect they will not be a problem since it is designed for my objective.
X.8 Sensitivity Analysis	Sensitivity analysis would be a great help. More specifically, sensitivity analysis to my inputs would help me identify if I can even make prediction based on the fidelity of my intelligence.	No sensitivity analysis is available. I can run multiple manual iterations to bound the problem.
X.9 Confidence Reporting	Confidence reporting would be great to have.	No confidence reporting is available.

Everything matches on the model, but the software has no sensitivity analysis capability. Sensitivity analysis would be useful to show where to expect diminishing returns in terms of man hours, or perhaps show the best way you and your wingman could bracket the widest possible values to increase the chance for success.

7. Combat Identification and Automatic Target Recognition (Confidence Reporting)

You are flying a wartime mission and have an unidentified contact that you want to engage. An automatic target recognition system, which combines data from several sensors, displays that the contact is an enemy. You push the pickle button and destroy the target. Are you sure that you just destroyed the enemy?

Table 8. Combat Identification (Confidence Reporting)

Areas of Interest	About Me	About the Model
X.1 Objective	Determine if the contact is an enemy.	Identify if the contact is an enemy.
X.2 Functional Area	I need to do analysis.	The model is built for analysis.
X.3 Scope	I need a subsystem level model.	It is a subsystem level model.
X.4 Deterministic or Stochastic	I need a deterministic model.	It is a deterministic model.
X.5 Data sources and Calculation	I need high accuracy.	The model relies on two different sensors and combines the information to determine result.
X.6 Descriptive or Prescriptive	I need at least a descriptive model, but it can be prescriptive as long as I can believe in it.	The model completes the identification matrix so that I can shoot and is therefore prescriptive.
X.7 Chief Abstractions	I can accept anything that does not affect the accuracy of the identification.	Before the sortie I did not identify the models abstractions.
X.8 Sensitivity Analysis	I would like to know if there is any possibility that a small change in one parameter would change the answer the model gives.	Sensitivity analysis is not available
X.9 Confidence Reporting	Confidence reporting has life or death importance.	No confidence reporting is available.

This example could occur in either an air-to-air or air-to-ground context, since the combat identification problem occurs in both arenas, but the following discussion will pursue an air-to-air example.

In the past, we relied on operators to provide the common sense approach to fusing what different sensors, i.e. models, were reporting, with their knowledge of the situation to derive how certain he was the result was accurate. If there was no way that a MiG could be at this location, then doubt entered the picture and sometimes, very correctly, the pilot would not shoot. Most performance measures for identification gave the percentage of time that the system gave the correct identification when it was looking at a given target type. Assuming that sensor A is looking at a MiG-29, it reports MiG-29 80% of the time. When flying a mission, the operator used those probabilities and situational awareness to decide if it made sense that this was an enemy contact.

Probability notation is useful in describing the intuitive weighting that pilots did. In probability notation, a line after the first condition means that the second condition is assumed. For example, $P("M"|M)$ would mean the probability (P) that the system shows MiG ("M"), given that the system is looking at a MiG ($|M$). Similarly, $P\left("O"|M\right)$ means the probability that the system shows other given that it is looking at a MiG. To find the probability that when your system said MiG you were actually looking at a MiG, you need to find $P(M|"M")$. Note that the order has changed. We now need to find the probability that we are looking at a MiG given that our sensor reports a MiG. The math relies on something called Bayes' Theorem and produces:

$$P(M|"M") = \frac{P("M"|M)P(M)}{P("M"|M)P(M) + P("M"|O)P(O)}$$

The only reason the equation is shown is to show that to find the desired probability, you multiply by the probability that you actually encounter a MiG. This is what the pilot intuitively did, when he weighed whether it was likely that there was a MiG here. Note also, that as the chance of running into other airplanes that could be misidentified goes up, seen as the term $P("M"|O)P(O)$, then the chance that you are rely looking at a MiG goes down.

When the pilot did intuitive weighting, the math involved is immaterial. For an ATR to do this, it will require built in logic. Whether the abstractions are valid will be the chief question. In our example, how will the automatic target recognition system combine the multiple systems feeding the algorithm? What sort of math are they doing? This is the first issue you must understand.

Just as important, how sure are the algorithms of the results? The key for both the operators and designers will be to ensure that the operators ultimately understand the outputs and have some way of weighing their accuracy. Confidence measures could include reliability ratings proposed in Section VI, or geometric shapes to convey system certainty. The integration must also preserve a method for the operator to drill down into the system and find what the individual sensors are saying. This could also prevent training to a capability that a resourceful enemy may defeat.

8. Campaign Planning (Applicability of Abstractions)

You are in a planning cell trying to decide the best manner to attack terrorist organizations. At your disposal, you have a software suite that is designed to identify the most lucrative targets for your effects based planning. It models the terrorists as "Dynamic Bayesian Networks." The software also claims that it will find the best course of action by picking the targets and designating the timing of actions and attacks. How do you treat the output from the software?

Table 9. Campaign Planning (Applicability of Abstractions)

Areas of Interest	About Me	About the Model
X.1 Objective	Find the most effective course of action.	Find the most effective course of action.
X.2 Functional Area	Analyze	Analyze
X.3 Scope	Campaign	Campaign
X.4 Deterministic or Stochastic	Deterministic	Deterministic
X.5 Data sources and Calculation	You require a relative assessment of the importance of each course of action.	The model provides the best course of action.
X.6 Descriptive or Prescriptive	Descriptive	Descriptive
X.7 Chief Abstractions	The chief abstractions cannot make the answer irrelevant.	To do its calculations the software relies on "Dynamic Bayesian Networks"
X.8 Sensitivity Analysis	I would like some way to gauge how critical given assumptions are.	There are tools available for manual sensitivity analysis.
X.9 Confidence Reporting	Since the software is deterministic, any measurement of confidence will rely on how good the assumptions are.	There is no formal way of reporting confidence. In this case my awareness comes from discussion with intelligence analysts about how good the assumptions are.

We will not delve into the specifics of this model, but rather the process one would need to follow. Your first challenge is to find out what are "Dynamic Bayesian Networks." Without this understanding, you will have no way to verify if this abstraction fits your problem. You learn the model centers on how one individual or group influences another. You believe this sort of network presentation is valid for this terrorist organization and run the model. The model identifies several targets that you had already considered, further reinforcing your belief of their validity. It also presents another target that you had not considered. You find this very useful, since after further consideration it may be valuable if intelligence can verify certain characteristics.

You then turn your consideration to the recommended solution that the model has put out. You disagree with the sequence. You recommend the solution you had originally conceived, and through an understanding of the model assumptions, can guess why it does not arrive at the same conclusion as you.

9. Campaign Planning Continued (Objective)

You are planning the campaign against an enemy nation and you are concerned about the enemy Integrated Air Defense System. You have a piece of software available that models battles, and someone suggests that you input the Air Tasking Order for the first day of the war and see what happens. In this particular case, there is no easy way to implement this solution, but the modelers begin jury-rigging a system to import the ATO data. After much labor someone recognizes that "predicting" the outcome of the battle is a waste of time and the effort is scrapped.

Table 10. Campaign Planning Continued (Objective)

Areas of Interest	About Me	About the Model
X.1 Objective	Predict if the plan will work.	Model relative merits of various courses of action in an air battle.
X.2 Functional Area	I need to do analysis.	The model is designed for analysis.
X.3 Scope	I need a battle level model.	It is a battle level model.
X.4 Deterministic or Stochastic	I want a stochastic model to include the element of chance.	It is a stochastic model.
X.5 Data sources and Calculation	Someone wanted an accurate prediction of the future.	The model uses individual entities that interact in a battle space to find the results. It does the calculations for the engagement using your starting parameters and data on each individual entity.
X.6 Descriptive or Prescriptive	I need a description of what will happen.	It is a descriptive model.
X.7 Chief Abstractions	Since the goal was to predict the future, I cannot accept abstractions that will affect that.	Each entity has its own data. The currency of this data is unknown.
X.8 Sensitivity Analysis	I would like to know how important given assumptions are.	Sensitivity analysis is not available.
X.9 Confidence Reporting	I do not consider confidence reporting.	Confidence reporting is not available.

The only time that models will prove predictive of the future is when they model a physical process that is well understood. For example, it is a possible to predict the phases of the moon based on knowledge of orbital mechanics. Yet with the higher level of this model, there are too many interactions and parameters to say that this model will "predict" the future.

This endeavor has validity only if the desire is to compare several different options and chose which is best. Say for example, that the plan could include stealth assets or not. How important is that contribution? Wary that it is a stochastic process, you consult an analyst who recommends a certain number of iterations for each scenario. You run both scenarios and compare the results. With stealth, your losses were half as large. Of course, this is not a measure of the real losses, but it does provide a strong indication of relative improvement when using stealth.

You also decide to check the data sources and calculations. You recognize that this model is driven by data from other models; it is essentially built on the assumptions of the smaller models. You wonder about how current and accurate the data is, but you are fortunate to have the time and assets available to verify that all entities involved in the model perform close to real life expectations (flagging would be helpful, see Section VI). You further examine how the model does its engagements and find that for the one-on-one engagements the values are very close. For the many on many results, the model provides a much more conservative answer, but by understanding the model you recognize that the model does not make allowances for synergistic cooperation of blue forces and feel that such red cooperation is unlikely.

VI. Takeaways for Modelers

Modeling literature discusses all the issues that I have covered. However, most of the operator audience has not been exposed to this literature. In the era of the PC, there are more and more cases of operators using model results without having had any indoctrination about their limitations. For this reason, it is critical that very basic information be included up front with all models that can potentially end up running as stand-alone components. Modelers can take steps to reduce the potential misuse of models or facilitate their use. Some models already have some or all of the mechanisms I will discuss, but there is no uniformity.

My proposition is that the questions posed in the matrix be included with the top level of help screens. The model must also flag its analysis capability. Finally, for an analysis model, modelers should incorporate sensitivity analysis and measures of confidence described in terms understandable by the operator. It is imperative that all this information be embedded with the model, not in associated briefings orF help files that may become divorced from the model.

As an example, I will use a fictitious missile evaluation software named Missile Fly-out Model. With the help drop down, there should be an option titled “Model Summary.”

Table 11. Example Model: Missile Fly-out Model 2.0

Areas of Interest	
X.1 Objective	Organization XYZ designed Missile Fly-out Model (MFM) 2.0 to provide a high fidelity prediction of air-to-air missile capability versus maneuvering and non-maneuvering targets across a wide range of altitudes and airspeeds.
X.2 Functional Area	MFM is designed for analysis.
X.3 Scope	This is a system level model that uses all known data to model motor, guidance logic and seeker performance.
X.4 Deterministic or Stochastic	MFM is a deterministic model. MFM has no capability to model the random characteristics such as pointing errors possible in seeker position, differences in missile battery life, separation effects, etc.
X.5 Data sources and Calculation	MFM uses tables of values for each missile that summarize key engineering data. See help under “missile parameters” for complete list of parameters. The user enters shooter and target parameters to define each situation. Target maneuvers can only be in relationship to distance from shooter and will only be to specified headings. A known issue is that target maneuver G are not limited by altitude, so users must specify logical values. The tabular and user input data is then used by the model to calculate each missile trajectory
X.6 Descriptive or Prescriptive	This model is descriptive only. Results for missile fly out assume that each missile is performing optimally and there is no randomness in fly-out termination. For example, if battery life is input as 20 seconds, all model replications will terminate at 20 seconds. See individual missile help files for real world variability information.
X.7 Chief Abstractions	MFM 2.0 has an ACF level of IV. The program itself uses a pseudo 5-degree-of-freedom (5-DOF) model to simulate aircraft and missile flight dynamics. The program treats missiles as a single point with x-, y-, and z- coordinates to describe its position. The model then finds an angle -of-attack and yaw angle for that position from a table of values. The code calculates a position change by applying velocities and accelerations over a finite increment of time -thus the "time-step simulation. The practical result is that for all heart of the envelope and long-range shots, the calculations will be extremely accurate. Only for extremely dynamic shots, such as maneuvering combat is there potential for errors. In this regime, effects such as wing twist, missile tip-off, etc. may become issues.
X.8 Sensitivity Analysis	MFM 2.0 has no capabilities for doing sensitivity analysis on missile parameters. MFM 3.0 will include the ability to vary tabular and user data over a variety of values.
X.9 Confidence Reporting	Because MFM is deterministic, the only confidence reporting is the flagging of the ACF for each missile. See help file “XX” for description of color codes. These codes are visible for all output screens as stoplights next to WEZ plots.

Modelers can summarize analysis capability for users with the proposed stop lights in Figure 3. This Analysis Capability Flag (ACF) is intended to summarize key issues for operators. Operators will have to note that the ACF values apply only for the objective for which the model is built.

Each model would have an ACF that is visible in the up-front summary. This overall ACF would pass information to the user about the where the model is in its life-cycle and how far down the road to absolute performance the analysis can proceed. Table 12. Analysis Capability Flag Definitions provides a description of these values.

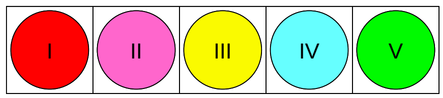

Figure 3. Analysis Capability Flag

Table 12. Analysis Capability Flag Definitions

Level	Meaning For Modelers	Meaning For Users	Example
I (Red)	This model is not designed for analysis and should not be used in that fashion without contacting model producer to see if analysis is possible.	Do not use this model for analysis.	A training flight simulator
II (Pink)	This is a new model whose chief abstractions have not been verified.	This model is new, use with caution.	A model designed to predict the effects of high-powered lasers, where lasers of sufficient power have not yet been developed to verify experimentally its outputs.
III (Yellow)	This is an intermediate value, whose significance will be judged by the model builder. It suggests that the model has reached a moderate level of refinement, but still has not fully matured.	The relevance to the user will depend on how the modeler uses this value.	
IV (Blue)	The chief abstractions in this model have a strong correlation to verifiable results. This is the highest level that battle and campaign levels can reach. For engineering level, it suggests further refinement above level IV.	For engineering level models, there are still some abstractions present. It is probably a simplification of another model. If greater fidelity is required the other model should be used. For all models above the engineering model, it implies that the model can only be used to weigh relative merit between options.	A battle model whose individual entities have been well documented.
V (Green)	This level is only attainable by system and sub-system levels since it equates to an absolute predictive capability. This model covers all pertinent factors and the calculations and tables involved have been verified with real world results.	This model can be used for prediction of absolute capabilities. The chief limitations present may be with individual entities within the model.	A fly-out model for a well understood missile.

For the models that have more than one entity, then each entity would have its own ACF, summarized within the help files and called for applicable outputs.

Table 13. Entity Analysis Capability Flag Values

Level	Meaning For Modelers	Meaning For Users	Example
I (Red)		This output should not be used	Single run of stochastic model
II (Pink)	New Model Entity	This entity is new and the parameters are highly speculative	Brand new air-to-air missile
III (Yellow)	This is an intermediate value, whose significance will be judged by the model builder. It suggests that the model has reached a moderate level of refinement, but still has not fully matured.	The relevance to the user will depend on how the modeler uses this value.	
IV (Blue)	The chief abstractions in this model have a strong correlation to verifiable results. There may subtle simplifications that may limit accuracy of output.	The relevance to the user will depend on how the modeler uses this value.	A thrust vectoring missile that can be approximated by algorithms, but has some limitations.
V (Green)	This model covers all pertinent factors and the calculations and tables involved have been verified.	This entity or output has the highest degree of accuracy possible and matches real world results	A missile or radar that has been extensively tested on range.

The utility of two-types of flagging is evident in the missile fly-out model. The software engine for the calculation relies on the laws of physics. In this particular case, there is a slight abstraction to simplify the math and the modeling of target maneuvers. Thus, the model would get a Level IV flag. Within the model, each individual missile would have its own color code, based on how good the data values for its tabular data are. Figure 4. Missile Comparison with Analysis Capability Flags, shows how two missiles, A and B could be compared and the different Analysis Capability Flags.

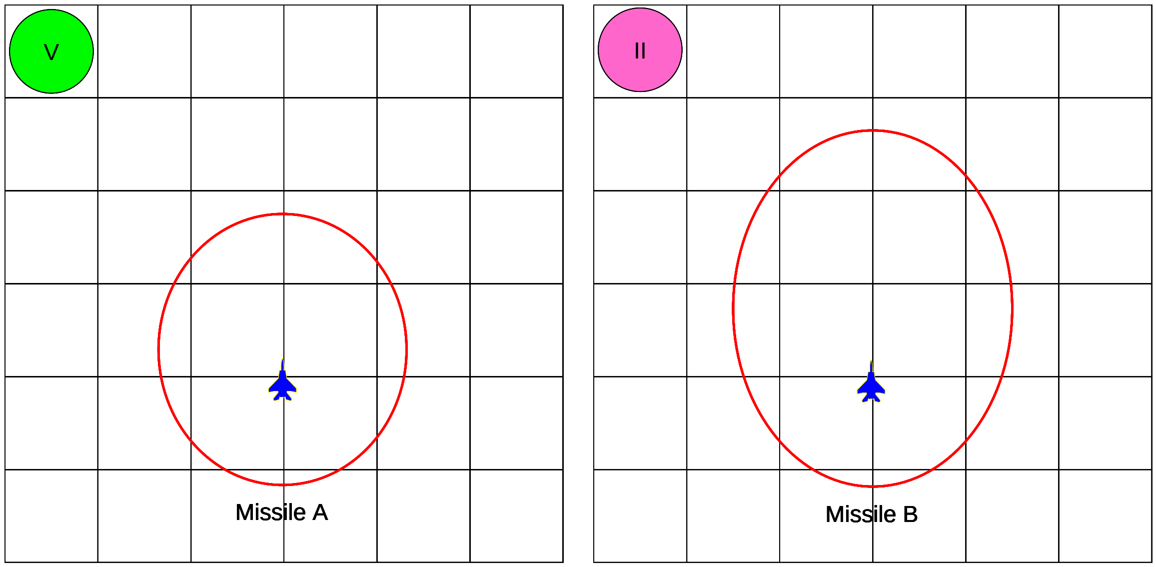

Figure 4. Missile Comparison with Analysis Capability Flags

Note that the ACF does not imply anything about the usefulness or value of models. It is simply a way of communicating known levels of abstraction and assumption to audiences. For example, modeling a brand new threat missile is incredibly valuable and provides a baseline amount of knowledge, but it would have a Level II flag. The purpose of this flag would be to warn users that there is a great deal of uncertainty about this model. The ACF value can also decrease to communicate that some new insights have shown limitations to the model.

In some ways, models are repositories of our institutional knowledge. Their importance to planning and operations demand customer management systems that pass the most current information to users and modelers who rely on given models. This discussion is beyond the scope of this paper. Analysis Capability Flags could serve as trips for automated customer management systems to push new information to appropriate customers. The term customer must also include the models on the higher levels of the cube. Those models rely on the outputs of the lower level, and so flagging information that should be pushed upwards could reduce the lag time in those models (Hughes, 1997: 47).

This ACF can also be used for confidence reporting. My description of confidence reporting has been very general and oriented towards the operator so far. The actually means of confidence reporting would be defined by the modeler, based on what is appropriate for the nature of the model. For example, a stochastic model could map confidence intervals to the confidence coding. Most of the times, operators are looking for simple displays of confidence. We need to know if the information is strong enough that we can do something with it. Low confidence information may still have value, especially if it is the only information available.

In my opinion, there is room for improvement in giving sensitivity analysis tools to users. For engineering models, like a missile fly out model, operators must manually run several iterations to get an idea of the importance of one input parameter. Simple built in sensitivity analysis tools could save many hours. Sensitivity analysis does not have to be limited to just user inputs, but can also include estimated parameters within the model to provide a bounding of outputs for the user.

Other key issues from an operator perspective are presenting reliability information and preserving the ability to "drill-down" in to menus. For example, an automatic threat recognition capability that synthesizes several inputs to identify a target is great. However, it must preserve some confidence estimate. In the heat of the battle, it gives the operator some idea of how reliable the information is which he can weigh into the decision. It must also preserve the ability to look at individual outputs. This ability to “drill down” is helpful, especially when one considers potential enemy countermeasures, stale data, or system degradations.

The final issue to consider was raised with question 3.6 in Section V. If the operator has output from a prescriptive model that does not agree with his own ideas, it creates a huge dilemma. If he does not follow the model, and the results turn out badly, others may hold that against him. Similarly, if he follows the model’s recommendation and it turns out badly he could be accused of incompetence. The root cause is a lack of transparency in the model. If the user cannot identify what the model is doing, it is hard to identify why the two disagree. Thus communicating the model’s methodology is critical for all involved (Fleishmann and Wallace: 2004).

This paper has spent a fair amount of time discussing how operators can make errors with models because they do not understand the assumptions. Modelers are vulnerable to the same issues, but in reverse. Every real world operation has specific objectives that drive the operation. Just as is the case with models, these objectives drive what are valid questions to ask about the exercise. Modelers attempting to construct models above the system level have the difficulty that they are often modeling activities with which they have no first hand experience. A logical source of data would seem to be real world operations, but the modeler can only draw data where there the real world objectives match the purpose of his model.

Red Flag is a good example of this potential data mining error. The objective for Red Flag is to provide a demanding simulation of the first handful of combat sorties for inexperienced aircrew. The goal is to create as difficult and demanding a situation as possible so that aircrew can reinforce those habit patterns that are most beneficial in combat situations. Using USAF Aggressor aircraft to fight the blue forces is a key element in creating this demanding situation. Statistics on exchange ratios between blue forces and the Aggressors simulating threats would seem to be a valid way of collecting data on performance. This overlooks the objective of the exercise. To keep a demanding environment, the Aggressors regenerate frequently. This regeneration is of great training value, but may or may not be representative of a realistic threat. Thus, an exchange ratio may be skewed by the objective to train. Similarly, the Aggressors will also vary their replication of pilot capabilities, meaning that each day of the exercise is not the same. Blue objectives must also be included. Sometimes blue objectives may drive identification criteria to force visual merges. The result is that it is very difficult to draw modeling data from such an environment.

This is an air-to-air example, but the same difficulties apply for air-to-ground exercises. For example, if a modeler is searching for fratricide information, it may seem like drawing information from joint exercises would seem logical. The joint exercise may be designed to exacerbate weaknesses in communications links to fully examine their impact. In this case, data on fratricide may be unreasonably high. The same type of issues could apply to target identification or time to find targets. It even applies to deployments and multi-national exercises, that may also have specific objectives that skew data.

The bottom line is that both communities must understand the objectives and limitations of the data or simulation are drawing from. In fact, my experience has been that operators most often learn about modeling issues when they work directly with the modelers. Conversely, modelers built the best models when they work closely with operators.

VII. Conclusions and Recommendations

Conclusions of Research

The farther away we get from killing the enemy when he fills up our windscreen, the more likely it is that we have to rely on a model to get the same kill. For that reason, we cannot afford to misunderstand and misapply models. The structure I propose is only a first step, but hopefully a logical one, in developing a broad baseline understanding of the new tools of our trade.

Significance of Research

This paper should make it easier for operators to ask the right questions to understand the models they use. If modelers can also convey those answers in the standardized format suggested, all parties would benefit. Operators would have one stop shopping to gain a basic understanding of their models, and modelers could feel more comfortable that they have communicated the chief limitations to the user audience.

Recommendations for Action

Implement procedures to educate model users about fundamental benefits and limitations of products. Implement a standardized approach to presenting information to users to make the education process more straightforward. Put another way, we must institutionalize transparency.

Operators should have a means to search and find models for their purposes. This central location should also include brief summaries of issues contained within the matrix. Similarly, a dedicated help line or access point for operators to access modeling experts would be beneficial.

Recommendations for Future Research

This paper approached the subject of models from my experience as a user. My perspective is colored by my training and by my past as an F-16 driver, but I hope that the approach is general enough that it can capture the process for all communities. If there is value to this approach, then verification of wording and the approach should occur with a cross functional, even cross service, committee.

Bibliography

Bard, Jonathan F. and Paul A. Jensen. *Operations Research Models and Methods*. John Wiley & Sons Inc., 2002.

Chamberlain, Tyrell. "Unguided Weapons Delivery Theory." USAF Weapons School Academics Course EMP300_04A, Barksdale AFB: 340th Weapons Squadron, 2004.

Clemens, Robert T. and Terence Reilly. *Making Hard Decisions with Decision Tools, 2001, 2nd rev.* Duxbury Thomson Learning.

Department of Defense (DoD). *Directive Number 5000.59.* Washington: Department of Defense, October 1995.

Devore, Jay L. *Probability and Statistics for Engineering and the Sciences with Infotrac*. Thomson Learning, 2003.

Fleishmann, Kenneth R. and William A. Wallace. "A Covenant with Transparency: Opening the Black Box of Models," *Communications*, 48: 93-97 (May 2005)

Friel, John. "Air Battle Models," *Military Modeling for Decision Making, 3 ed*. The Military Operations Research Society, Inc., 1997.

Hartman, James K. *Lecture Notes in Aggregated Combat Modeling*. Ed. Jack Kloeber Jr and Jack Jackson. 1985.

Hartman, James K. *Lecture Notes in High Resolution Combat Modeling*, 1985.

Hughes, Wayne P. Jr. ed. *Military Modeling for* Decision Making, 3 ed. The Military Operations Research Society, Inc. 1997.

Kelton, W. D. and others. *Simulation with Arena*, *3 ed*. McGraw-Hill, 2004.

Law, A. M., and W. D. Kelton. *Simulation Modeling and Analysis, 3 ed*., McGraw-Hill, 2000.

Miller, John O. Class slides, OPER 677, Modeling and Analysis of Air Operations. Department of Operational Sciences, Air Force Institute of Technology, Wright-Patterson AFB OH, Spring 2005.

Ross, Sheldon M. *Introduction to Probability Models, 8 ed*. Academic Press, 2003

Wentz, Larry K. and Lee W. Wagenhals. "Effects Based Operations for Transnational Terrorist Organizations: Assessing Alternative Courses of Action to Mitigate Terrorist Threats." San Diego CA: 2004 Command and Control Research and Technical Symposium, 2004.

Additional Sources

Betts, William D. *Conventional Attack Planning with CWDS 9.1*. Student Paper, USAF Weapons School F-16 Class WSF 02 BIF. USAF Weapons School, Nellis AFB NV, December 2002.

Bigelow, James H., William W. Taylor, S. Craig Moore, Brent Thomas. *Models of Operational Training in Fighter Squadrons.* Santa Monica CA: RAND, 2003.

Bosone, John W. *Safe Escape in the F-16*. Student Paper, USAF Weapons School F-16 Class AIF. USAF Weapons School, Nellis AFB NV, 14 June 2003.

Caroli, Joseph, Daniel Fayette, Nancy Koziarz, and Terrance Stedman. "Tools for Effects Based Course of Action Development and Assessment." 2004 Command and Control Research and Technical Symposium, San Diego CA. 2004.

Cook, Bryan E. *Calculating Release Altitudes for Results*. Student Paper, USAF Weapons School F-16 Class 98 BIF. USAF Weapons School, Nellis AFB NV, 12 December 1998.

Dana, Matthew R. *All Weather PGMs*. Student Paper, USAF Weapons School F-16 Class 98-AIF. USAF Weapons School, Nellis AFB NV, 13 Jun 1998.

Davis, Paul K. "Exploratory Analysis and Implications for Modeling," *New Challenges, New Tools for Defense Decisionmaking*, Santa Monica: RAND, 2003.

Fantini, Michael A. *F-16 Attack Planning for Training and the "Real Thing*." Student Paper, USAF Weapons School F-16 Class 93 BIF. USAF Weapons School, Nellis AFB NV, 18 Dec 1993.

Falzon, Lucia and Jayson Priest. *The Centre of Gravity Network Effects Tool: Probabilistic Modelling for Operational Planning.* Edinburgh South Australia: DSTO Information Sciences Laboratory, 2004.

Fogle, James C. *The Block 40/42 Navigation Solution*. Student Paper, USAF Weapons School F-16 Class 93 BIF. USAF Weapons School, Nellis AFB NV, 18 December 1993.

Hicks, J. Todd. *Viper's Guide to Combat Weapon Delivery Software (CWDS).* Student Paper, USAF Weapons School F-16 Class 96 AIF. USAF Weapons School, Nellis AFB NV, June 1996.

Lesman, Michael F. *Deploying with Situational Awareness Data Link* Student Paper, USAF Weapons School F-16 Class 02 AIF. USAF Weapons School, Nellis AFB NV, June 2002.

Malone, Timothy G. and Reagan E Schaupp. "The 'Red Team': Forging a Well-Conceived Contingency Plan," *Aerospace Power Journal*, AU Press: 16: 22-33 (Summer 2002)

Murphey, Todd A. *F-16 JDAM Accuracy vs. Terminal Threats*. Student Paper, USAF Weapons School F-16 Class 99 AIF. USAF Weapons School, Nellis AFB NV, 6 June 1999.

Pirnie, Bruce, *Analysis of Special Operations Forces in Decision Aids: Recommendations*. Santa Monica CA: RAND, 1994.

Pitts, Curtis. *Combat Programs and Products for the 21st Century Weapons Shop*. Student Paper, USAF Weapons School F-16 Class 03B. USAF Weapons School, Nellis AFB NV, 13 December 2003.

Podolak, J. P. *WCMD Planning and Employment Considerations for the F-16.* Student Paper, USAF Weapons School F-16 Class 02 BIF. USAF Weapons School, Nellis AFB NV, December 2002.

Schroeder, Jeffrey C. *CWDS 7.0: Mission Planning for the Nintendo Generation*. Student Paper, USAF Weapons School F-16 Class 98 BIF. USAF Weapons School, Nellis AFB NV, 12 Dec 1998.

Secretary of the Air Force. *Air Force Policy Directive 16-10: Modeling and Simulation Management.* Washington: Secretary of the Air Force, 30 January 1995.

Wagenhals, Lee W. "An Operational Concept for Effects Based Course-of-Action Development and Evaluation in a War Game." *Enabling Technologies for Simulation Science VI*, Alex F. Sisti, Dawn A. Trevisani, Editors, Proceedings of SPIE Vol. 4716 2002

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 074-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
03-06-2005	Graduate Research Project	Jun 2004 – Jun 2005

4. TITLE AND SUBTITLE A PRIMER AND GUIDE TO MODELING FOR OPERATORS	**5a. CONTRACT NUMBER** **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Jansons, Juris L., Major, USAF	**5d. PROJECT NUMBER** **5e. TASK NUMBER** **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 640 WPAFB OH 45433-8865	**8. PERFORMING ORGANIZATION REPORT NUMBER** AFIT/GOS/ENS/05-09
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSOR/MONITOR'S ACRONYM(S)** **11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Operators rely more and more on models to accomplish their work; examples include the weapons employment zone displays in cockpits, logistics models for deployment, and battle simulations to decide courses of action. They often do not have much exposure to modeling, and the products they are using do not always supply adequate documentation. The first portion of this paper serves as a primer on modeling for operators. It then proposes a matrix of questions that an operator should know to ask about any model he is using. The next section contains several examples to illustrate the discussion. The last section includes a proposal to use the matrix as a standard format for modelers to pass relevant information to users. If the operators know which questions to ask, and modelers can embed that information inside the models, then overall effectiveness should increase.

15. SUBJECT TERMS

Modeling and Simulation.

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Miller, John O. ENS
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)* (937) 255-6565, ext. 4326
U	U	U	UU	83	(john.miller@afit.edu)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18